ES.1803: Symmetric Matrices

Jeremy Orloff

1 Introduction

Symmetric matices are very important in math, science and engineering. Our ultimate goal is to prove the following theorem.

Spectral Theorem. A real $n \times n$ symmetric matrix has n orthogonal eigenvectors with real eigenvalues.

The generalization of this theorem to infinite dimensions is widely used in math and science. In fact, the Fourier series that we study in ES.1803 can be seen as an application of this theory.

2 Symmetric Matrices

We can understand symmetric matrices better if we discuss them in terms of their properties, instead of their coordinates. To avoid being too abstract, we will rely on coordinates for the following two definitions.

Definition 1. For column vectors $\mathbf{v}$, $\mathbf{w}$ the inner product is defined in terms of transpose and matrix multiplication: $\langle \mathbf{v},\, \mathbf{w} \rangle = \mathbf{v}^T \mathbf{w}$.

(In 18.02 you called this the *dot product*.)

Definition 2. The matrix A is symmetric if $A^T = A$.

Property 1. Suppose A in an $m \times n$ matrix and B is $n \times p$. Then, $(AB)^T = B^T A^T$.

Proof. Let $A_{i,j}$ be the i, j entry of A. Writing out matrix multiplication in terms of indices:

$$\left((AB)^T\right)_{i,j} = (AB)_{j,i} = \sum_{k=1}^{n} A_{j,k} B_{k,i} = \sum_{k=1}^{n} (B^T)_{i,k} (A^T)_{k,j} = (B^T A^T)_{i,j}. \qquad \blacksquare$$

2.1 Properties of symmetric matrices

Property 2. If A is symmetric, then $\langle A\mathbf{v},\, \mathbf{w} \rangle = \langle \mathbf{v},\, A\mathbf{w} \rangle$.

Proof: We use the definition of inner product:

$$\langle A\mathbf{v},\, \mathbf{w} \rangle = (A\mathbf{v})^T \mathbf{w} = \mathbf{v}^T A^T \mathbf{w} = \mathbf{v}^T A \mathbf{w} = \langle \mathbf{v},\, A\mathbf{w} \rangle.$$

(The second to last equality follows because $A = A^T$.)

Everything we do below will follow from this Property 2.

Property 3. If A is symmetric and $\mathbf{v}, \mathbf{w}$ are eigenvectors with different eigenvalues, then $\langle \mathbf{v},\, \mathbf{w} \rangle = 0$, i.e., $\mathbf{v}$ and $\mathbf{w}$ are orthogonal.

Proof: Suppose $A\mathbf{v} = \lambda_1\mathbf{v}$ and $A\mathbf{w} = \lambda_2\mathbf{w}$. Then using Property 2, we see

$$\lambda_1\langle \mathbf{v},\, \mathbf{w}\rangle = \langle \lambda_1\mathbf{v},\, \mathbf{w}\rangle = \langle A\mathbf{v},\, \mathbf{w}\rangle = \langle \mathbf{v},\, A\mathbf{w}\rangle = \langle \mathbf{v},\, \lambda_2\mathbf{w}\rangle = \lambda_2\langle \mathbf{v},\, \mathbf{w}\rangle.$$

Look at the first and last terms of this chain and use $\lambda_1 \neq \lambda_2$ to conclude $\langle \mathbf{v},\, \mathbf{w}\rangle = 0$. ■

Property 4. If A is real and symmetric, then A has a real eigenvalues.

Proof. We'll give both an algebraic and analytic proof of this key property.

Algebraic proof: If we had defined the term, I could just wave the word Hermitian. As it is, I will do the same thing without using the word. For non-zero complex vectors, note two things:

1. $\langle \mathbf{v},\, \overline{\mathbf{v}}\rangle = \sum \mathbf{v}_k\overline{\mathbf{v}}_k > 0$.
2. Since A is real, $\overline{A\mathbf{v}} = A\overline{\mathbf{v}}$.

If $\mathbf{v}$ is an eigenvector with eigenvalue λ, then

$$\lambda\langle \mathbf{v},\, \overline{\mathbf{v}}\rangle = \langle \lambda\mathbf{v},\, \overline{\mathbf{v}}\rangle = \langle A\mathbf{v},\, \overline{\mathbf{v}}\rangle = \langle \mathbf{v},\, A\overline{\mathbf{v}}\rangle = \langle \mathbf{v},\, \overline{A\mathbf{v}}\rangle = \langle \mathbf{v},\, \overline{\lambda\mathbf{v}}\rangle = \overline{\lambda}\langle \mathbf{v},\, \overline{\mathbf{v}}\rangle.$$

Looking at the first and last terms in this chain, we see $\lambda = \overline{\lambda}$. This proves λ is real. ■

Analytic proof: This proof is a litlle more involved than the algebaic one, but it produces an eigenvector and some geometric insight.

Let S be the unit sphere, i.e., the set of all vectors of length 1.

We can use Lagrange multipliers, with the objective function $f(\mathbf{v})$ and constraint $g(\mathbf{v}) = \langle \mathbf{v}\,\mathbf{v}\rangle = 1$, to find the maximum of f on S.

The symmetry of A implies $\text{grad} f = 2A\mathbf{v}$ and $\text{grad} g = 2\mathbf{v}$. So the Lagrange multiplier equations are

$$2A\mathbf{v} = 2\lambda\mathbf{v}, \quad g(\mathbf{v}) = \langle \mathbf{v},\, \mathbf{v}\rangle = 1.$$

The first equation shows that constrained critical points occur when $\mathbf{v}$ is an eigenvector. Therefore, the point $\mathbf{v_1}$ on S where f has a maximum must be an eigenvector. ■

3 Spectral Theorem

Spectral Theorem. Suppose the $n \times n$ matrix A is symmetric. Then it has n orthogonal (hence independent) eigenvectors with real eigenvalues.

Proof: Property 4 provides one eigenvector, $\mathbf{v}_1$ with eigenvector λ_1.

Let W be the set of all vectors orthogonal to $\mathbf{v}_1$. First we show that if $\mathbf{w} \in W$ then $A\mathbf{w} \in W$. To do this we must show $\langle \mathbf{v}_1, A\mathbf{w}\rangle = 0$ for all $\mathbf{w} \in W$:

$$\langle \mathbf{v}_1, A\mathbf{w}\rangle = \langle A\mathbf{v}_1, \mathbf{w}\rangle = \lambda_1\langle \mathbf{v}_1, \mathbf{w}\rangle = 0.$$

But now the same argument as in Property 4 (replacing $\mathbf{R}^n$ with W) shows that A has a real eigenvector $\mathbf{v}_2 \in W$. We can continue this until the eigenvectors form a basis of $\mathbf{R}^n$.